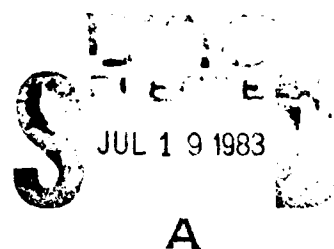Compilation
of
Cooperative Data Element Dictionary
of
Five Federal Agencies' Systems
for
Processing of Technical Report Literature

to
U.S. Department of Commerce
National Technical Information Service

Madeline M. Henderson
Contractor

1 March 1983

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ii

EXECUTIVE SUMMARY

The data element dictionary project consisted of the compilation of a dictionary encompassing all of the data elements used in the major systems for handling technical reports; the systems involved are those of the Department of Defense Technical Information Center (DTIC), Department of Energy Technical Information Center (DOE/TIC), the National Aeronautics and Space Administration Scientific and Technical Information Facility (NASA/STIF), the National Technical Information Service (NTIS), and the Government Printing Office (GPO). All those elements, both substantive and "housekeeping" types, used in processing technical reports in each system are included in the compilation; each record contains the element's definition, rules for use, the tags or indicators used internally, and those used externally for information interchange.

The work of selecting the data elements and collecting the information about them was accomplished by the individual agencies; GPO offered the use of its computer facility for the compilation and for preparation of indexes to the individual dictionaries and to the combined lists.

The data element dictionary (DED) is a useful tool for the participating agencies, constituting current documentation of their systems, probably in greater detail than heretofore available. In addition, the DED can serve as a tool for the agencies to examine the desirability and feasibility of further standardization to improve cooperative processing among them. Such an effort is probably most immediately attractive for the four agencies with COSATI-based indexing systems: DTIC, DOE/TIC, NASA/STIF, and NTIS. Their systems are variations on a set of principles embodied in conventions of long standing.

However, the maximum value of the DED can lie in its use as a mechanism for those agencies and their library-type, MARC-based cataloging counterpart, GPO, to explore productive avenues of cooperation and interfacing. The five agencies all process Federal technical reports, obviously duplicating efforts and producing separate and

iii

(presently) incompatible services in the form of data bases and printed bulletins, catalogs, and indexes. Considerable gains in efficiency, elimination of redundancy, and improved timeliness in availability of bibliographic records can result from shared input processing of those documents.

It is recommended here that the DED form the basis for consideration by the COSATI-based systems of a single acceptable set of common (core) data elements with rules and external tags agreed upon by the four agencies. Such a set of standardized data elements should be convertible to each of the current systems, with a minimum of cost to each. Furthermore, such a set of standardized elements can be compared to the set of MARC-based system elements on a 1:1 basis.

This latter comparison, it is further recommended, can facilitate effective cooperation between the COSATI-system agencies and the GPO. A proposal has been made, and generally accepted, for experimentation with an intermediate working file of technical reports' minimum-level processing records. The data elements, rules for their use, and tags identifying them need to be defined and accepted by the five agencies. The data elements should be those sufficient to identify a document but not all those necessary for any one agency to discharge its responsibilities for control of documents and production of services based on bibliographic records. In such an experiment, the agreed-upon rules will permit an agency to enter the core data elements into a working file; when other agencies receive the same document, they can retrieve the record and convert it to their own formats. Thus agencies can have a means for control and processing of reports records in a more efficient and timely manner.

I. Introduction

A. Background of Project

This project was an outgrowth of an earlier study of coopera-
tion in the processing of the technical report literature. (1) That
study involved detailed comparison by the principal investigator of
the available documentation for four processing systems: that of the
Defense Technical Information Center (DTIC), of the National Aero-
nautics and Space Administration Scientific and Technical Information
Facility (NASA/STIF), of the Department of Energy Technical Informa-
tion Center (DOE/TIC), and of the National Technical Information
Service (NTIS).

The report of the study recommended the compilation of a dic-
tionary of the data elements (DED) used by the four systems, to
include definitions and rules for use of the elements. As recom-
mended, the compilation was to be based on input prepared by the per-
sonnel actually involved with operation of the systems; the principal
investigator would expedite the compilation but the agencies would be
responsible for selection of elements and collection of information
about them.

As noted in the report of the earlier study, that effort was
needed because of the "proliferation of data bases, the growth in num-
ber of users of information systems, technological advances and devel-
opments in resource sharing [which] all contribute to concern about
lowering the costs and increasing the speed and efficiency of Federal
reports-processing operations." The concerns also included the feasi-
bility of formulating a network concept, the need for modifications of
existing systems and the impacts and benefits of any modifications or
adjustments for each agency and for the user community.

----

(1) Study of Cooperation in the Processing of the Technical Report
Literature to the U.S. Department of Commerce, National Technical
Information Service, Madeline M. Henderson, Contractor, 13 August
1980.

Objectives of the earlier study therefore included determining
the need for and expected benefits of development of additional stan-
dards, or update of standards already in place, to improve the coopera
tive processing of technical reports; and the desirability of and
potential for closer coordination between such standardization efforts
of the technical reports community, on the one hand, and those of the
library community and its networking activities, on the other.

These objectives, it was noted in the study report, differed
from earlier proposals for the effort in the emphasis on relations
between the technical reports and library communities. The two inter-
face with each other and with users at a number of points; coordinat-
ing those interfaces can contribute to improving cost-effectiveness
and efficiency of Federal bibliographic data processing.

B. Additional Related Activities

The report of the earlier study referenced several related
activities within the library community. These included a review of
the MARC formats for cataloging library materials and efforts to
develop a minimum-level or simplified cataloging record. The report
noted the relation between these activities and those of the techni-
cal-reports processing community as represented by the study project
itself.

For example, during the conduct of this project, the principal
investigator participated with Libary of Congress (LC) Staff and Gov-
ernment Printing Office (GPO) catalogers in identifying which data
elements should be mandatory and which optional for national-level
bibliographic records, both full records and minimum-level records,
for technical reports. Late in 1981 the decision was made to wait foɪ
completion of the DED before undertaking further efforts to so desig-
nate the data elements.

Such mutual interests converged in a discussion meeting which took place on February 12, 1981 and involved representatives of NTIS, the Federal Library Committee (FLC), the user community, and the GPO. The discussion focused on possibilities for further cooperation in processing (cataloging) Federal documents and technical reports. GPO had been considering merging GPO and NTIS bibliographic data files to improve user access to information about government documents; this discussion explored the concept of interfacing the files, rather than merging them. The proposal was made, and endorsed, to explore the feasibility of developing a model working file for shared input processing of Federal technical reports. Several real problems were noted, however, including identification of a computer system which could accommodate the working file. But as a first step, it was agreed, a data element dictionary should be compiled, to include not only the elements of the four major reports-processing agencies but also those used by GPO in MARC-based cataloging of reports.

Later in 1981, NASA/STIF and DOE/TIC entered into an agreement, later joined by DTIC, to explore the benefits of an inter-agency information and data exchange program involving the "sharing of bibliographic and numeric data bases through the intercommunication" of the agencies' ADP equipment. The agreement was an extension of efforts at Lawrence Livermore National Laboratory to provide a network interconnecting DOE computer sites for data exchange. The so-called Intelligent Gateway Computer System is the mechanism for accomplishing automated access procedures for various data bases such as DOE/RECON and NASA/RECON, and post-processing capabilities such as extraction and format specification.

As such improved access and post-processing capabilities become available, it is obvious that users will benefit from more standardized data elements common to a number of data bases. The DED effort will serve as a basic tool to assist the agencies in working toward further standardization of at least their core data elements.

Another effort under way during this period involves the Resource Sharing Advisory Group to DTIC and its Committee on Cataloging Rules (RSAG/CCR). The group is working over the COSATI cataloging rules with a view to coordinate and further standardize the way in which the cooperating agencies (DTIC, NTIS, and contractor laboratories) catalog technical reports they all handle. Here again, the DED will constitute a basic tool for comparison of definitions and current rules for applying specific data elements.

A companion effort exists in the library community: MARBI is an American Library Association joint committttee concerned with representation in machine readable form of bibliographic information; its responsibilities include encouraging the creation of needed standards and continually reviewing existing standards. MARBI deliberations impact on MARC formats, such as are used by GPO in cataloging technical reports, and are therefore of interest to this project and the DED.

C. Objectives and Tasks

Based on the results of the earlier study, and cognizant of additional related activities such as those noted, a proposal was made for the cooperative development of a data element dictionary for Federal documents. It was noted therein that the purpose of such a DED is two-fold: first, to provide to the reports-processing agencies a tool to guide their consideration of possible further standardization to achieve greater compatibilities and improved cooperative processing among themselves; and secondly, to provide to them and their library counterparts a mechanism by which to explore productive avenues of cooperation and interfacing.

The original proposal made a distinction between the data elements used by GPO in its library type of cataloging, and the data elements proposed by the Library of Congress (LC) for the cataloging of

- 4 -

technical reports within the family of MARC formats. In actual fact, the two sets of data elements, although then in process of final selection and definition, were one and the same; therefore, the DED has five sets of data elements, not six.

The objective, it was proposed, would be met by a series of tasks involving the principal investigator and personnel of the participating agencies. The tasks, to be accomplished in sequence, were:

1. To define computer requirements for the compilation, in the format recommended as most promising, in terms of computer processing and final layout.

2. To negotiate for use of an appropriate computer system, hopefully "donated" for the project because of its perceived value.

3. To develop the instructions for inputting data elements, based on suggestions from the participating agencies and requirements of the donated computer system.

4. To supervise the routine input of data elements, based on the developments of Task 3.

5. To prepare the dictionary in final form, to oversee the final print-out, and to prepare a final project report.

II. Conduct of Project

The project was initiated in July 1981 with the issuance of a contract by NTIS; what had been projected as a 3-month to (perhaps) 6-month effort actually took closer to 18 months. A calendar of key dates is included as Appendix A; the more detailed discussion which follows will put the events into focus.

A. Tasks 1 and 2

In September 1981, at a meeting of representatives of management for the participating agencies, a plea was made for computer support. Approximately two weeks later word came that GPO would support

the project with its computer facilities. After submitting a preliminary description of the project, the principal investigator met with key ADP personnel at GPO early in November to describe the project and its computer requirements in more detail. As a result, a systems analyst, A. D. Lowry, was assigned to work with the principal investigator in defining requirements in a more structured manner.

A meeting of agencies' staff involved with cataloging/indexing operations was held in November, at which time the data elements to be recorded were agreed upon: all the elements used for processing technical reports, including "house-keeping" type elements but not including data elements needed for processing books or journal articles. In addition, the rules for recording descriptions for each data element were drawn up; a copy is attached as Appendix B. These rules constitute a refinement and extension of earlier rules drafted by the principal investigator, and benefited from the input of the agencies' participants.

The work in defining computer system requirements proved to be a valuable effort: although somewhat time-consuming, the result was a well-developed statement of functional requirements. This is attached as Appendix C; it can serve as a guide to other similar efforts to compile a DED.

Once the GPO ADP facility had accepted the statement of system requirements, it issued a tape format specification which triggered participating agencies' preparation of input data.

B. Tasks 3 and 4

The task of defining the data elements and rules for inputting them was accomplished in November 1981, as noted above. When the GPO tape specifications were distributed, the agencies started to input their own data. As could be expected, the agencies with the most complete and current documentation of their systems were first to input tapes to GPO; DTIC and DOE both had their material in in April 1982.

- 6 -

NTIS staff developed their input data more slowly and experienced some difficulties with tape preparation; however, its final good copy was sent to GPO in November 1982.

GPO itself developed its descriptions in conjunction with LC; Library staff prepared the definitions and rules for use for each data element, then GPO staff confirmed their use in actual cataloging operations. Several data elements therefore carry the note "GPO not presently using pending decision." The GPO data were input via terminal to the ADP facility; the terminal used accepted upper-case characters only, hence the difference in appearance of the GPO dictionary.

The NASA/STIF is operated under contract, so considerable negotiation between NASA headquarters management and contractor staff was necessary to effect a viable project. As a result, NASA was last in (December 1982) with its data.

The role of the principal investigator during this period, obviously much longer than anticipated, was as a resource in solving problems of selection or definition of data elements, as a bridge between the GPO computer facility staff and the agencies' participants, and as a source of encouragement to those participants.

C. Analysis of Dictionaries

In June 1982, the managers of the five agencies met with the principal investigator to discuss progress and possible further activities. It was possible to give each person a copy of the then-existing DED, namely, the dictionaries of DTIC and DOE plus their respective indexes and the consolidated indexes. In addition, DTIC had prepared a KWIC index of the element names.

Reviewing the contents of this abbreviated DED, it was pos-
sible to note the similarities and differences apparent in the data
elements, and their rules for use. Based on the KWIC index, for
example, DTIC and DOE seem to have the following data elements in
common:

abstract	descriptors/posting term
corporate author	distribution (in some form)
personal author	note (of one kind or another)
availability (in some form)	number (accession/file/identifica-
classification	tion)
corporate/source code	pages/pagination
contract/grant number	price/NTIS price
date/report date	subject codes/categories
	title (in some form or another)
	type of item

This list is similar, but not identical, to the list of common
data elements included in the report of the earlier study. The differ-
ence can be explained by the fact that the latter list was of elements
common to all four systems studied, not of only the two, DTIC and
DOE/TIC.

Subsequent to the June meeting, and as additional pieces of
the dictionary became available, the principal investigator has had
the opportunity to examine the preliminary listings in more detail. A
matrix of all elements in all systems is shown in Table 1.

It became apparent that any list of common data elements needs
to be reviewed by the agencies to determine if these elements are actu-
ally in common, and whether they are all the common elements. (Some
element names obscure similarities, others suggest commonality which
doesn't exist.) A listing of seemingly common data elements for all
systems is shown in Table 2.

TABLE 1

Matrix of Data Elements and Systems

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
Accession/Identification Number	x	x	x	x	x
Subject Fields and Groups	x	x	x	x	
Corporate Author	x	x	x	x	x
Corporate Code		x		x	
Title	x	x	x	x	x
Corporate Source Supplement			x		
Classified Title			x		
Classification of Title	x		x		
Descriptive Note	x		x	x	x
Personal Author(s)	x	x	x	x	x
Report Date	x	x	x	x	x
Pagination (page count)	x	x	x	x	
Source Series/Report Number	x	x	x	x	x
Secondary Report Number		x			
Contract/Grant Number	x	x	x	x	x
Project Number	x		x	x	x
Task Number	x		x	x	x
Monitor/Sponsor Acronym	x		x	x	
Monitor/Sponsor Series	x		x	x	
Classification (of Report)	x	x	x	x	
Supplementary Note/Drop Note	x	x	x	x	x
Distribution (Availability) Statement	x	x	x	x	x
Subject Descriptors, Splits		x			
Descriptors	x	x	x	x	x
Subject Descriptor Count		x			
Classification of Terms	x				
Posting Terms/Identifiers	x			x	
Classification of Identifiers	x				

- 9 -

TABLE 1 (cont.)

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
File Number		x			
Descriptor Upposted Terms		x			
Abstract	x	x	x	x	x
Abstract Qualification			x		
Classification of Abstract	x				
Inventory (number of hard copies)	x			x	
Annotation (title modification)	x	x	x	x	x
Special codes (about availability)	x				
Reclassification Code	x				
Distribution (availability) Codes	x		x	x	
Special Codes/Control Codes	x			x	
Serial Number/Source Series	x		x	x	
Source Code	x		x		
Document Location	x			x	
DTIC Remote Terminal Identification	x				
Classification Authority	x				
Declassification Date and Event	x		x		
Downgrading Date and Event	x		x		
Geopolitical Code/Location	x				x
Organization Type Code	x				
Classification Extension Authority	x				
Review Date	x				
Classification Extension Reason Code	x				
Type of Handling/File Selected for (i.e., where data go)		x	x		
Issue Number (for announcement)			x	x	
Receipt Type (how acquired)			x		
Item Description (set)				x	

TABLE 1 (cont.)

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
Contract Numbers (all cited? Y or N)			x		
Highlight Notes				x	
Receipt Date			x		
Authors' Names (all cited? Y or N)			x		
Distribution Control/Foreign Source of Data			x		
Authority for Classification Field	x				
IAC Document Type and Code	x				
IAC Report Number	x				
IAC Subject Terms	x				
Country of Finance			x		
Language Code		x	x		x
Country of Origin/Affiliation		x	x		x
Copyrighted (Y or N)			x		
Copyright Override			x		
Reproduction Authorized (Y or N)			x		
Microfiche Code			x		
Microfiche Availability Code			x		
Microfiche Source Origin			x		
Hardcopy Availability Code			x		
Document Class (type) Code		x	x		
Abstract Source Code			x		
References (Y or N)			x		
Abstract's Author Identification			x		
Analytic Note (in lieu of abstract)			x		
Analytic Primary/Number of Subsidiaries			x		

TABLE 1 (cont.)

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
Analytic Subsidiary Subject Code			x		
Analytic Subsidiary Number			x		
Form of Document			x		
Language Note			x		
Author's Affiliation (other than Corporate Source)		x	x		
Author's Role (e.g., editor, compiler)			x		
Place of Publication (other than Corporate Source)		x	x		x
Publisher (other than Corporate Source)			x		x
Sales Agency and Pricing/Stock Number	x	x	x	x	x
Secondary Corporate Sources Code			x		
Special Publication Notes			x		
Available Supplements (e.g., films)			x		
Supplementary Information re Translation			x		
Unclassified Foreign Title			x		
Reprint Note			x		
Number of Volumes in Set			x		
Literary Indicator		x			
Subdistribution Category		x			
Source of Bibliographic Information		x			x
Other Announcements			x		
TIC Utilization		x			
TIC Identification Number		x			

- 12 -

TABLE 1 (cont.)

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
Conference Title		x	x		x
Conference Place		x			
Conference Date		x			
Report Origin (DOE project? Y or N)		x			
INIS Proposed Descriptors		x			
INIS Temporary Record Number		x			
INIS Document Type		x			
INIS Categories		x			
NTIS Note		x	x		
Leader, bytes 0-23					x
Logical record length					x
Record status					x
Legend					x
Type of record					x
Bibliographic level					x
Indicator count					x
Subfield code count					x
Base address of data					x
Encoding level					x
Descriptive cataloging form					x
Entry map					x
Record directory					x
Control number (OCLC)					x
Date and time of latest transaction					x
Reduction ratio (microforms)					x
Fixed length data elements					x
Date entered on file					x
Type of publication date code					x
International standard book number					x

TABLE 1 (cont.)

Data Element Name	DTIS	DOE	NASA	NTIS	GPO
Country of publication...code					x
Illustration code					x
Intellectual level code					x
Form of reproduction code					x
Nature of contents code					x
Government publication code					x
Conference publication indicator					x
Festschrift indicator					x
Index indicator					x
Main entry in body of entry indicator					x
Fiction indicator					x
Biography code					x
Modified record code					x
Cataloging source code					x
LC card number					x
Standard technical report number					x
Coded mathematical data					x
Cataloging source					x
Geographic area codes					x
LC call number					x
Copy, issue, offprint statement					x
Geographic classification code					x
NLM call number					x
NAL call number					x
GPO item number					x
Dewey Decimal Classification number					x
Edition statement					x
Mathematical data area					x

TABLE 1 (cont.)

Data Element Name	DTIC	DOE	NASA	NTIS	GPO
Physical description (AACR II)					x
Series statement - title (traced)					x
Series untraced or traced differently					x
"With" note					x
Dissertation note					x
Bibliography/discography note					x
Contents note (formatted)					x
Photoreproduction note					x
Local note					x

Note 1. Some data elements are necessary for inventory or distribution control; some are required for the production of printed services, i.e., abstract/index bulletins.

Note 2: For the sake of clarity and ease of comparison, the principal investigator combined some data elements in some systems to correspond to single elements in others, and recorded like elements under one name, in spite of differences in names among systems. The marks in the cells of the matrix represent the principal investigator's interpretation of the contents of the individual DED's; these judgments will need to be examined by the individual agencies.

TABLE 2

List of Common Data Elements

Accession Number
Corporate Author
Title
Personal Author(s)
Report Date
Source Series/Report Number
Contract/Grant Number
Supplementary Note
Distribution (Availability) Statement and Date
Subject Descriptors
Abstract
Title Annotation (title modification)
Sales Agency and Pricing/Stock Number

When the common data elements have been identified, the rules for recording them need to be studied, to determine the extent of differences and whether such differences are necessary. For example, for the data element "abstract," DTIC rules allow for 1800 characters of narrative text ("about 200 words"), NTIS says 3500 characters maximum (and "usually limited to 200 words"), NASA limits the text to 150 words, and DOE calls for a maximum of 2,000 characters; all allow free-form text with alphanumeric and special characters. The GPO data element "Summary, abstract, scope, etc. note" provides for unformatted notes which can constitute a summary or abstract; no mention is made of any arbitrary length limitation. In general it would appear that the five systems have essentially the same rules for this common data element.

But the data element "personal author" presents a different picture. DOE calls for a standard format: last name, a comma, a space, an initial, a period, an initial, a period. Additional authors are entered in the same format, separated by a semicolon and a space; no mention is made of a limit to the number of authors' names allowed. Diacritical marks are omitted, according to DOE rules. But DTIC says "up to 5 entries," with 60 characters for the first and middle name or initial, and 60 characters for the last name, which is preceded by a / indicator, "The names...are entered as they are recorded on...the document...as complete a name as possible." Alphanumeric and special characters are allowed.

GPO provides for recording personal names as main entries or added entries, in accordance with AACR II cataloging rules. The personal name entry "may consist of words, phrases, initials...;" AACR II allows diacritical marks and calls for the "most commonly known" name by which a person is known or "most commonly found" form of the person's name. The name is entered by the element of the name "under which the person would normally be listed"; this is generally therefore the surname. (There are some 47 pages of rules in Chapter 22, "Headings for Persons" in AACR II, 2nd Edition.) The personal author or principal personal author is entered as the main entry; in

case of shared personal authorship, the first person named is entered as the main entry and additional authors' names are entered as added entries. (This rule probably has more importance for technical report cataloging than for cataloging literary works.)

In the NASA system, the inverted name is preceded by a precode number, then entered: last name, first name, middle name or initial; a maximum of 10 authors is recorded. "Succeeding authors are separated with an @ sign." NTIS, on the other hand, divides the personal author field into five subfields, each with a 50-character limit. In each subfield, the first name and/or initials is followed by a "reverse slash" and the last name. The subfields are separated by commas; the first name is preceded by the word "by" and the last name in the list is preceded by "and."

This range of rules and conventions suggests the scope of the effort that may face either the reports-processing agencies alone, or them and their cataloging counterparts at GPO in concert. When the rules are converted to sample entries, the similarities and, especially, differences become clearer:

        DOE : Cardoon, G.B.
        DTIC: George Bernard/Cardoon
        GPO : Cardoon, George Bernard
        NASA: 01 Cardoon, George Bernard
        NTIS: by George Bernard/Cardoon

These differences can be meaningful to the user: are G.B.Cardoon and George Bernard Cardoon the same person? Possibly, but how about entries for J.P. Jones, John Paul Jones, and James Peter Jones: will the differences cause searching problems? These questions have more significance, perhaps, in the context of the Intelligent Gateway Computer System development described earlier, and in terms of improved coordination between reports indexing and library cataloging procedures.

Identifying the core data elements common to the five systems is one step in determining how best to standardize the expression of the information they each and all process. Exchange of data among the reports indexing systems is currently accomplished by transmission of bibliographic records, primarily on magnetic tape and formatted according to community standards (e.g., the American National Standards Institute Standard format, ANSI Z.39-2). This standard includes identification of specific data elements by means of 3-digit tags. The DED descriptions of data elements includes explication of those tags, as assigned by each of the systems for each data element.

If the exchange and sharing of bibliographic data are to be improved for those agencies responsible for processing and control of Federal technical reports and similar documents, it seems reasonable to examine the tags used to identify common data elements. The closer the agencies can come to common identification, the closer they can come to efficient exchange.

But tags assigned to the core data elements differ: they are all 3-digit tags, but the same 3 digits are used for different elements in the systems. For example, DOE uses the tag 110 for its data element Primary Title; DTIC uses the same 3 digits to identify the element Report Classification; in the GPO system, 110 identifies Main Entry - Corporate name; NASA doesn't currently use the tag 110; and for NTIS, it is the tag for the data element Classification (NTIS and DTIC are closer in their use of various conventions than any other agencies). The external tags for common data elements are given in Table 3.

Now the use of 3-digit tags is well defined in each system, but as the systems share bibliographic data between and among themselves, the differences in tag use can cause complications, if not problems. Each agency must accommodate its processing steps to the differences in tags, translating incoming data to its own requirements. For example, NTIS must convert the DOE tag 110 to its own tag 220 (Title) or perhaps 270 (Title Annotation); the DTIC tag 220 (Unclassified Title) is supposedly converted to 270, but the NASA tag 145 is equated to the NTIS 220. Though procedures have been established to accomplish

TABLE 3

EXTERNAL TAGS FOR COMMON DATA ELEMENTS

	DTIC	DOE	NASA	NTIS	GOP
Accession Number	001	290	---	001	074
Corporate Author	300	710	142 143	300	Main Entry, 110 Subj. Add. Entry 610 Add. Entry,710 Series Add. Entry,810
Title	220	110	145	220	Main Entry,130 Uniform Title,240 Subj. Add. Entry,630 Add. Entry,730 Series Add. Entry,830
Personal Author	280	070	150	280	Main Entry,100 Subj. Add. Entry,600 Add. Entry,700
Report Date	480	370	095	480	008
Report Number(s)	170	Primary, 150 Secondary, 210	185	170	088 STRN, 027
Contract/Grant Number	320 330	240	179	320 330	536
Supplementary/Drop Note	580	440	148	580	740

Exhibit 3 (cont.)

	DTIC	DOE	NASA	NTIS	GOP
Distribution/Availability Statement	130	510	--	130	506
Subject Descriptors	380 400	General, 801 Splits, 802 Count, 800	Major, 197 Minor, 198	380 400	Subj. Add. Entries 600, 610, 611, 630, 650, 651, 653
Contract	620	950	249	620	520
Title Annotation/Augmentation	270 540	620	087 160	270 540	245
Sales Agency and Pricing	530	430	191	530	037

this sort of translation automatically, the questions can still be raised: are the differences necessary? can the agencies agree on a common set of tags for their common (core) data elements for purposes of exchange and sharing of bibliographic data? It must be remembered that the tags are used to facilitate sharing, and should not affect the internal operational systems unique to each agency. But sharing and exchange are important, for increased efficiency in the processing of Federal technical documents.

A further complication is introduced when considering the tags assigned by GPO, according to MARC cataloging rules, for the data elements it shares with the indexing systems of the other four agencies. Thus GPO makes a distinction between Main Entry-Uniform Title Heading (Tag 130), Subject Added Entry-Uniform Title Heading (Tag 630), and Added Entry-Uniform Title Heading (Tag 730). The conventions are admirably consistent within the MARC system, but don't "match" the tags assigned by the other agencies. Again, there is no need for or interest in coordination or standardization, except for considerations of sharing bibliographic data to improve productivity, red·· ʒ duplication of effort, and ensure more timely identification of Federal technical data and information for i_ʃ effici·e·at use.

As noted earlier, the examination ʃf the data elements, their definitions and rules for use, and the consideration ʃf possible coordination or standardization of those rules, need to be done by representatives of the agencies participating in the overall effort. Those involved with procedures, responsibilities, missions, operational requirements, and all the facts of daily existence can better judge the advantages and disadvantages of changes looking toward community coordination and improved compatibility. The few examples of analysis of the DED conducted and reported here serve to illustrate the utility of the DED as a tool to facilitate the kind of dialog which is both desirable and possible, given current pressures and technological opportunities.

D. Task 5

The last task proposed for the project was compilation of the
DED in final form and preparation of an accompanying report. The form
of the DED as distributed to the participating agencies provides a
working tool, a form that permits examination, tracing, comparison,
reference, and careful diagnosis as recommended in the proposal: a
tool for the reports-processing agencies to guide their consideration
of possible further standardization among themselves, and for those
agencies and their library-cataloging counterpart to explore coopera-
tion and interfacing between them.

In its current "final" form, the DED will probably not prove
as effective in apprising users of the various systems about the con-
tent of those systems: the DED needs further work of an editorial
nature to improve its utility outside the participating agencies. The
principal investigator has agreed with the agencies that they need a
chance to review their individual data in light of all input, in order
to arrive at more nearly uniform definitions, rules, and descriptions
throughout the compilation. Such uniformity can improve the utility
of the DED for users of the various data bases and reference products.
However, the distribution of the DED describing current systems should
not wait on consideration of possible modifications leading to further
standardization among the systems. Users are faced with the current
operational systems and can benefit from timely availability of the
DED for those systems.

E. Reports to the Community

In the course of this project, a number of requests came from
various segments of the library and information-service communities
for reports on the progress of the effort and for participation in
related programs. A major example of the latter was the examination
of data elements for national-level bibliographic records for techni-
cal reports, described earlier in this report.

Early in 1981, before the project actually got underway, the principal investigator prepared an article for Science and Technology Libraries, in which the interest in improved standardization and increased cooperation among Federal agencies responsible for processing technical reports was described. (2)

Even earlier, another article was written on standardization in general but describing the interests of Federal agencies in improving reports processing through increased standardization. (3)

Other opportunities to report to the community were presented at various meetings and conferences. In September 1981, at the meeting of the Depository Library Council to the Public Printer, the principal investigator described the purpose, scope, and aims of the project to the Committee on Bibliographic Control of the Council.

A workshop sponsored by the National Bureau of Standards' Federal Information Processing Standards activities (NBS FIPS) on the subject of Software Documentation provided another forum for discussion of the project. This presentation was part of a session on Enhancing Software Sharing; the principal investigator stressed the use of the DED as a management tool and as one step in the process of improving resource sharing. (4)

---

(2) Henderson, Madeline M. Some Aspects of Technical Report Processing by Federal Agencies, Science and Technology Libraries, 1:4, 1981 Summer: 19-26.

(3) Henderson, Madeline M., Standards Developments and Impacts, Special Libraries, 72.2:2; April 1981: 142-148.

(4) Henderson, Madeline M., Compilation of Bibliographic Data Element Dictionaries, in NBS FIPS Software Documentation, Proceedings of a Workshop held March 3, 1982 at NBS, Gaithersburg, MD. NBS Special Publication 500-94, Washington, DC, U.S. Government Printing Office, October 1982: 209-214.

Two different forums for discussion of the project were pre-
sented in April 1982, at meetings of MARBI and the RSAG/Cataloging
Rules Committee. Both of these groups and their goals were described
earlier in this report. At the MARBI meeting at the Library of Con-
gress, the principal investigator's status report was part of a more
general discussion of minimum level cataloging, series/name authority
records, and particular aspects of MARC formats.

The RSAG/CRC meeting was devoted to more specific discussions
of recommendations for change to the COSATI-based cataloging guide-
lines followed by the four reports-processing agencies. The principal
investigator stressed the value of the DED as a tool to support deci-
sions about modifying standards to improve cataloging operations.
Such use of the DED will prove even more valuable for considering
changes across four systems rather than the one (DTIC's) of concern to
the RSAG.

An update on progress in compiling the DED was presented at
the annual American Library Association meeting in July 1982, to the
Government Documents Round Table Federal Documents Task Force (GODORT
FDTF). In particular, the Acquisitions and Bibliographic Control Work
Group had arranged a meeting with several reports of interest, includ-
ing the one on the DED.

Later in July the principal investigator joined colleagues at
the Naval Research Laboratory (NRL) Management Information Center to
discuss mapping of data in COSATI records to MARC format files. NRL
is one of several users of the Integrated Library System (ILS) devel-
oped at the National Library of Medicine who are interested in "hand-
ling both open literature and technical reports using ILS software."
ILS accepts records in the MARC format; the topic for discussion was
the possibility of redefining the MARC format to accept COSATI catalog-
ing records. The DED again could be used as a tool to support deci-
sions of this kind, and especially when such decisions affect more
than one of the major reports-processing systems.

At the invitation of DOE/TIC management, the principal investi-
gator visited the Oak Ridge offices to discuss in more detail the
goals of the DED project and its expected importance to individual
agencies and the community as a whole. The concept of working towards
shared bibliographic processing while not affecting individual sys-
tems' operations was a principal topic during the visit. There had
been opportunities to hold similar discussions with the other agencies
since they are local; the Oak Ridge visit rounded out the process of
communication.

The most recent report to colleagues was made in November 1982
at a meeting of the Committee on Information Hang-Ups. Since this
group initiated the early discussions which resulted in the DED pro-
ject, the members are especially interestd in its progress. Committee
members have concerns such as those expressed at the NRL meeting: the
need to handle both library cataloging records and technical reports
indexing records and the wish to handle them with the same or compat-
ible systems. Again the objectives of the project, to provide a tool
for standardization decisions and also for users' information, were
stressed. The other reports at this Committee meeting were related:
an NRL update on compatibility in local automation and a status report
on the Intelligent Gateway Computer System.

III. Uses of the DED

A. Agencies' Reviews.

As has been noted several times, the DED can serve the individ-
ual agencies who have contributed to it as a tool for their internal
use. The data element descriptions document current systems, in some
cases in more detail than has been available up to now. The exercise
of identifying and defining the elements in current use is helpful to
system managers and operating personnel. Further, the chance to read
how the other agencies define their elements, especially the elements
common to all the systems, is of interest purely from the editorial
standpoint. It has been interesting to read the texts for a given
data element as prepared by five different people:

sometimes just the wording obscures the fundamental similarities in definition, rules for use, and conventions applied to the element.

It is possible that each agency will opt for editorial revisions to improve clarity and uniformity of expression, especially for publication for the user community.

If no further use is made of the DED, the process of compiling it has been useful. A machine-readable file of data elements and their descriptions would be a useful product for the participating agencies; such a file could be maintained and updated more easily than most other forms of system documentation.

B. Comparisons of Data Elements

The potential kinds of analysis that can be derived from the DED have already been suggested: comparisons of data elements themselves as they differ from one system to another, of the rules set forth for the same or similar elements in the different systems, and of the tags or indicators used to identify the elements in exchange procedures. It must be recognized, however, that any one of these kinds of comparisons could be a lengthy process and one of questionable value, if carried to extremes.

A reasonable approach would seem to be to concentrate on the data elements common to a number of the systems. The core elements common to all five systems have already been identified; the list needs to be reviewed by system personnel, as suggested above. An initial effort might concentrate on the four COSATI-based systems, and involve their personnel in trying to reach agreement on their common data elements. The systems have the advantage of all being based on indexing principles and roughly adhering to a set of descriptive cataloging rules of long standing. Focusing on the degree of similarity existing among the various formats, the effort should result in recommendations for enhancing the common features and contributing to more efficient and less costly information exchange. In addition, the work might result in a single embodiment of the variations on the basic

indexing system, and therefore a better interface with the MARC-based cataloging system.

Then the list of elements common to the five systems, whatever its final form, can serve as the basis for discussions between the reports-processing agencies (DOE, DTIC, NASA, and NTIS), and their library-cataloging counterpart, GPO. The exercise would look to developing agreement on the essential data elements needed to identify a document in all five systems; on the rules for expressing those data elements in a manner useful for all five systems, not necessarily exactly as any system requires but so all systems can accept a bibliographic record and modify it to match internal operating requirements; and the tags/indicators to be adopted for such an interim recording system. The effort, if preceded by agreement among the COSATI systems, would mean comparing COSATI rules and MARC conventions on a 1:1 basis, to enhance possibilities of reaching agreement on rules for an intermediate system.

Considerable effort has been devoted to reaching partial understanding and agreements along these same lines: the work of the RSAG/CCR has already been noted, as well as that at the NRL information center. Members of RSAG/CCR have published an excellent report on their comparison of DTIC COSATI cataloging and MARC AACR II cataloging. (5) NRL staff continue to give status reports on their efforts also. However, both of these studies concern embodiments of one or two COSATI-based systems, not all four with their variations and differences. And they seek to provide for accommodation of the two approaches, COSATI indexing and MARC cataloging, in a single system, particulary the ILS mentioned earlier. In these respects, the efforts differ from the approach suggested in this project.

----

(5) Burris, Elaine, Sarah Mikel, Betty Pringle, Asta Kane, et al., Comparison of DTIC COSATI Cataloging and AACR-II Cataloging in the MARC Communications Format, RSAG/CRC-82/01, August 1982, 40p.

C. Value to Systems' Users

The DED can be of interest and value to the users of the various systems described through the dictionary. This was one of the early reasons for requesting development of a DED by the Committee on Information Hang-Ups. As noted in the earlier report (reference 1), the Working Group on Updating COSATI of the Committee discussed, as early as December 1978, the possibility of developing a data element dictionary as one step in investigating needs for better standards and protocols so that COSATI systems might "move toward consistent interaction with other national...networks..." and might "interact with other systems, such as MARC."

The DED was seen as a means for assisting in making information systems more compatible. Users accessing the various information systems available, especially through online computer-based data bases but also through printed indexes, catalogs, and abstract bulletins, face a number of problems due to differences among the systems. The problems and differences start with the rules for obtaining the access, continue through variations in formats and conventions for presenting bibliographic records, and reach to the contents of the records themselves, i.e., the data elements therein and the differences in recording those elements.

The DED for the five systems contained in it lays out all the data elements and the rules for recording them. The user can consult the dictionary to assist in developing search strategies for maximum efficiency in trying to identify documents, and can compare the contents of the individual files in order to select the most pertinent sources. This use of the tool is more important as more flexible access is provided by networks, developments such as the Intelligent Gateway Computer System, and similar technological advances.

One of the results of such developments and advances is that libraries and information centers, and those responsible for their management and the patrons they serve, are learning the value of access to a variety of sources, the value of resource sharing with each other, and the value of standards as a means to bring the variety and the sharing to pass.

IV. Conclusion

The effort to develop a DED has been a successful cooperative
endeavor on the part of the five agencies who have defined their data
elements in a structured way, perhaps more thoroughly than has been
done before.

Possible further activities can range from maintenance of the dic-
tionary in machine-readable form for ease of updating and distribu-
tion, to publication of studies describing similarities and differ-
ences among the various systems, to use as a tool with which to
develop new standards and improve existing standards for the biblio-
graphic data processing procedures of the cooperating agencies. The
principal investigator suggests that the next use of the DED should be
as a guide to the reports-processing agencies in considering possible
further standardization to achieve greater compatibilities and
improved cooperative processing among themselves. Another early use
should be to provide to those agencies and their library counterparts
a mechanism by which to explore productive avenues of cooperation and
interfacing. The DED will reveal, in both cases, the extent of common
ground in the identification and definition of data elements and the
rules for their use, as well as the magnitude of differences.

Beyond those immediate uses, what has been proposed, and accept-
ed as a reasonable experiment to undertake, is the development of an
intermediate working file to facilitate shared input processing/shared
cataloging. Participating agencies would enter a bibliographic record
containing the data elements identifying the particular document, fol-
lowing agreed-upon rules and using the specific tags established for
the file. When other participants receive the same document, they can
retrieve the record from the intermediate file and convert it (by
appropriate machine processing insofar as possible) to the format
required by the others' systems, either to enter into a data base or
to produce printed catalogs and announcement bulletins.

Such a shared-processing intermediate working file can eliminate
redundancy in analysis of Federal technical reports and speed up iden-
tification and total processing times.

APPENDIX A

Calendar of Key Dates

CALENDAR OF KEY DATES
Cooperative DED Project

March 11, 1981   : proposal submitted to NTIS by project principal
                        investigator

July 7, 1981   : contract issued by NTIS to principal investiga-
                        tor

August 1981   : dictionary compilation system documentation
                        received from consultant

September 25, 1981   : meeting of management representatives of
                        participating agencies on project goals and
                        procedures

Early October 1981   : offer by GPO to supply computer facility for
                        DED compilation

November 6, 1981   : meeting with cataloging operations personnel
                        from participating agencies to draft inputting
                        rules

November 9, 1981   : meeting with GPO ADP management staff to
                        discuss computer requirements

November-December
     1981   : work with GPO systems analyst to define system
                        requirements in General Functional Systems
                        Requirements (GFSR) document

December 11, 1981   : circulation of draft GFSR document to agencies
                        for comment

January-February
     1982   : revisions to GFSR document

February 23, 1981   : final version of GFSR document accepted at GPO
                        and distributed to participating agencies

March 8, 1982   : specifications for tape input issued by GPO to
                        participants

April 12, 1982	:	status report memo by principal investigator to participants
April 20, 1982	:	DTIC input tape ready for GPO
April 21, 1982	:	DOE input tape sent from Oak Ridge
May 17, 1982	:	no-cost extensio:. to original contract issued by NTIS
June 4, 1982	:	meeting of managers of participating agencies with principal investigator to discuss status and possible further activities
November-December 1982	:	GPO data input to computer facility
November 3, 1982	:	NTIS input tapes ready for GPO
December 10, 1982	:	NASA input tapes ready for GPO
January 20, 1983	:	first copies of completed DED made available to principal investigator and NTIS
January-February 1983	:	copies of DED delivered to participating agencies, with their input tapes
March 7, 1983	:	final report submitted to NTIS

APPENDIX B

Rules for Input to Date Element Dictionary Records

RULES FOR INPUT TO DATA ELEMENT DICTIONARY RECORDS

Record descriptors for each data element in the following order and as directed:

1. __NAME OF ELEMENT:__ Record the name of the data element as shown in your agency system documentation. Spell out abbreviations and add qualifiers, if necessary, to identify fully the specific data element.

   Input in free text form, alphanumeric characters, variable length, but not to exceed 200 characters (the equivalent of one 120-character line and one 80 character line).

2. __SOURCE AGENCY:__ Identify your agency, in whose system the data element appears and which is supplying the record.

   Select appropriate four-character acronym: DTIC, NTIS, NASA, DOES, GPOS.

3. __USE:__ Indicate whether the element is in one of four possible categories of use (called "validation" in DTIC system): mandatory, optional, mandatory if available, or required because of the presence of another cataloging/indexing entry.

   Select appropriate entry, 12 characters maximum: Mandatory, Optional, If Available, Required.

4. __DEFINITION:__ Record the narrative definition of the data element as shown in your agency system documenation.

   Input in free text form, alphanumeric characters, variable length, but not to exceed 720 characters (the equivalent of six 120-character lines).

5. __CONTENT:__ Indicate what is to be found in this data element in cataloging/indexing use. Include rules for recording the data element, and special codes to be used (if applicable). If this data element in a cataloging/indexing record consists soley of a code or abbreviation from a short list of alphanumeric codes or abbreviations, the list itself may be recorded. Indicate also any limitations imposed by the system such as field length limit, number of occurrences limit, etc.

   Input in free form, alphanumeric characters, variable length, but not to exceed 2400 characters (the equivalent of twenty 120-character lines).

6. __INTERNAL TAG:__ Record the exact content designation (often a mnemonic) by which the data element is identified for input procedures.

   Input up to 4 alphanumeric characters, left justified, with blanks if necessary.

7. INTERNAL TAG DESCRIPTION: Record any relevant indicators of subfield codes to accompany the Internal Tag (especially applicable to GPO data elements).

   Input in free text form, alphanumeric characters, variable length, but not to exceed 1200 characters (the equivalent of ten 120-character lines). If no description is available or relevant, input N/A (Not Applicable).

8. EXTERNAL TAG: Record the exact content designation by which the data element is identified in magnetic tape formats, as for exchange of bibliographic data among computer systems.

   Input three numeric characters.

9. EXTERNAL TAG DESCRIPTION: Record any relevant indicators of subfield codes to accompany the External Tag (especially applicable to GPO data elements).

   Input in free text form, alphanumeric characters, variable length but not to exceed 1200 characters (the equivalent of ten 120-character lines). If no description is available or relevant, input N/A (Not Applicable).

10. CHARACTER SET: Record a description of the possible characters that may appear for the given data element in a cataloging/indexing record in your specific agency system. Usual entries will be: alphabetic, numeric, alphanumeric, one of xx possible alphabetic codes, expanded 8-bit ASCII, etc.

    Input in free text form, alphanumeric characters, variable length, but not to exceed 50 characters.

11. DATA SOURCE: Indicate whether this data element in cataloging/indexing use is found in the source document, is created by the agency staff, is generated by algorithm or other automatic means, and whether it must be standardized to the agency system specifications whatever its source (e.g., corporate source code selected by agency staff).

    Input one or two of four entries, 8 characters maximum per entry (up to 16 characters maximum for descriptor): Document, Staff, Machine, Standard.

12. NOTES: Record any additional relevant information about the data element or its use in the specific agency system.

    Input in free text form, alphanumeric characters, variable length, but not to exceed 1200 characters (equivalent to ten 120-character lines).

APPENDIX C

General Functional System Requirements
for a Proposed
Data Element Dictionary (DED)

GENERAL FUNCTIONAL SYSTEM REQUIREMENTS

FOR A PROPOSED

DATA ELEMENT DICTIONARY (DED)

A COOPERATIVE PROJECT OF

THE UNITED STATES GOVERNMENT PRINTING OFFICE

THE NATIONAL TECHNICAL INFORMATION SERVICE

THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

THE DEFENSE TECHNICAL INFORMATION CENTER

THE DEPARTMENT OF ENERGY TECHNICAL INFORMATION CENTER

REVISED FEBRUARY 23, 1982

TABLE OF CONTENTS

I. Introduction

This paper describes the general requirement for a proposed bibliographic
data element dictionary (DED), encompassing the elements used in cataloging
and indexing Federal publications.

The purpose of this paper is to provide a general understanding of the objectives
for the DED, the overall parameters of the proposed system for generating
the DED, an estimate of expected processing volumes and frequencies, and
some fundamental assumptions and constraints.

To assist in the provision of a general understanding of the project, a series
of exhibits is attached to the narrative description: data element description
forms for the 12 descriptive items which will be captured for each bibliographic
data element; an input data source description and two sample input forms;
and 4 output product descriptions with 10 sample products.

Any questions or suggestions regarding this document or the DED project may
be addressed to:

> Madeline M. Henderson
> Consultant
> 5021 Alta Vista Road
> Bethesda, MD 20814
> (301) 530-6478
>
> or
>
> Albert D. Lowry
> Program Analyst
> Systems and Programs Division (SASP)
> U. S. Government Printing Office
> (202) 275-3269

II. Background

· Recent developments in the computer processing of bibliographic data have
included increased cooperation in inputting that data and more extensive on-line
access to the resulting files. These developments have been supported by
promulgation and use of appropriate standards which improve the chances for
cooperative processing and effective file access.

A recent study of bibliographic data processing systems of four major Federal
technical reports-processing agencies: Department of Defense (DTIC), Department
of Energy (DOE), National Aeronautics and Space Administration (NASA), and
National Technical Information Service (NTIS), recommended that the agencies
review their data elements to determine whether further standardization
should be undertaken. The study further recommended that the review should
include the data elements used by the Government Printing Office (GPO) in
cataloging Federal documents into the OCLC, Inc., computer center in Columbus,
Ohio.

To support this review process, the study recommended the compilation of a
dictionary of the bibliographic data elements used in the five systems,
including definitions and rules for applying the data elements, plus the tags or
indicators that accompany the data elements in use. The dictionary should take
the form of five separate compilations, each with its own indexes of element
names and tags, and with overall indexes of element names and tags.

The separate compilations will have immediate value to those involved in
management and operation of the individual systems, as well as to users of the
current systems. For example, GPO is in the process of updating its cataloging
system, and NASA has undertaken a review of its data entry procedures.

The DED's will prove to be useful tools in these internal projects. The
combined dictionary will allow the managers of the four reports-processing
systems to determine whether further standardization of their common data
elements will enhance compatibility and improve cooperative processing
among themselves. Further, the combined dictionary will provide to managers
of the five systems a means by which to explore productive avenues of
cooperation and coordination in the bibliographic processing of all Federal
documents. The dictionary will reveal the extent of common ground in the
identification and definition of data elements and rules for their use, as
well as the magnitude of differences.

The five agencies have supported the concept of compiling the dictionary of
their respective bibliographic data elements. NTIS is supporting the
principal investigator, M.M. Henderson, through a personal services contract;
all the agencies are providing staff time to identify and define their own
data elements and will prepare the input according to rules defined by the
investigator. In addition, GPO is examining the possibility of using its
computer facility to produce the compilation and indexes. DTIC has also
examined its computing capabilities and determined that it could assist in the
effort, as appropriate.

2

III. System Description

A. Objectives

The primary objectives of this project are to compile a dictionary of bibliographic data elements and to provide indexes for the effective use of the dictionary by the five participating agencies. Use of a computer facility will make the initial compilation faster, easier, more efficient, and more flexible. In additon, each agency will have a machine-readable file of its own bibliographic data elements for any further processing integral to its own operations.

To accomplish the objectives, input of bibliographic data elements will come from each agency for its own system. The input can be made—in a variety of ways, but the two methods seen as most reasonable for the project are by magnetic tape or on paper forms. In either case, the agencies will identify their own elements and record the required 12 descriptive items defined by the investigator for the project.

The computer facility will accept the input so as to format the dictionary compilation and print out the individual dictionaries with their own indexes, as well as the combined indexes to the total compilation.

B. General System Parameters

Data input to the computer will take the form of records containing 12 descriptive items about each bibliographic data element. The individual agencies will identify their own data elements and record for each: the element name, the agency identification, whether the element's use in cataloging or indexing is mandatory or not, the definition of the element, the content of the data element when used (i.e., the rules for using the data element, including any limitations on length or number of occurrences in the system record), the tag used in the system for input, description of that tag, the tag used in machine-readable exchange formats, description of that tag, the possible character set used for the element, whether the element is found in the document being cataloged or indexed, and any miscellaneous notes about the data element. These descriptive items are described in greater detail in terms of the computer facility's requirements in Attachment A.

The records, one for each data element from each agency and containing the 12 descriptive items, will be submitted to the computer facility in a format suitable for further processing. For example, the descriptive items which will be used for indexes (the name of the data element and the two types of tags) will be appropriately identified so that the computer can accomplish the necessary processing steps. The records will either be recorded on magnetic tape, according to specifications of the computer facility, or will be recorded on hard copy (paper) forms, for input by the facility itself. Further details about the input, including a sample paper form, are included as Attachment B.

The computer facility will accept the input tapes or enter data from the worksheets, to build five separate files, one for each agency. The data will probably come in from the agencies in order of use within their own systems, usually in order of input tags (numeric, alphabetic, alphanumeric, or some combination of the three). The computer facility need only build e files in the sequence of records as they are received. The computer facility will make available to the principal investigator interim printouts (unformatted "dumps")

3

for the purpose of proofing and editing as necessary. The computer facility will make the changes and corrections requested by the investigator. (The investigator will work with staff at the individual agencies to correct recurring errors in the content of their records, so as to minimize the burden on the computer facility.)

When the five individual files have been completed, the computer facility will process each one to produce the required indexes, then will process the total compilation so as to produce the overall indexes requested. The most desirable type of index of names of data elements is the keyword-in-context or permuted title index (so-called KWIC), in which all the major words in the names of data elements are brought to the left margin for alphabetic sorting. If a KWIC index is not feasible (i.e., if the computer facility does not have an operational program to produce KWIC indexes), then an alphabetized list of element names will have to do. The indexes by tags will be in order of numeric, alphabetic, and alphanumeric tags.

The end products which result from the computer processing will number 23 in all: five separate dictionaries (although they may be bound into one volume), each with its own name, internal tag, and external tag indexes, plus name, internal tag, and external tag indexes to the overall dictionary. Examples of these products are included in Attachment C.

C. Processing Volumes and Frequencies

1. Input Transactions

The five systems vary in the number of their data elements, from a low of perhaps 20 to a high of about 100. An absolute maximum number of records from all five systems will be 500; 250 may be closer to the number of data element records actually input. The investigator can only estimate the number at this time, since the participating agencies are in the process of identifying the data elements pertinent to processing reports and Federal documents in their own systems.

The size of each record also cannot be estimated accurately because of the variable length textual descriptive items (fields) to be included in each record (definition, rules for use, miscellaneous notes). Examples of input records included in Attachment B can indicate the probable size of individual records but no estimate of total volume can be made yet.

The input will be a one-time (possibly even all-at-one-time) transaction from each agency. If the decision is made to have the participating agencies input hardcopy forms to the computer facility for onsite data entry, the worksheets would be collected by the investigator, reviewed, and edited if necessary, then turned over as one batch for input.

2. Output Products

Except for the regular printouts for review and correction by the principal investigator, there will be one set of products from the total effort as described above. A total of perhaps 12 copies (2 for each agency and 2 for the principal investigator) will be requested of the total package of products.

4

D. Fundamental Assumptions and Constraints

1. The selection of data elements to be input, and the preparation of complete descriptions of each data element, are the responsibilities of the individual participating agencies.

2. The preparation of descriptions of the data elements will be accomplished, in tape or paper form, according to specifications developed by the computer facility.

3. Use of hardware and software in the computer facility will be at the discretion and under the supervision of the computer facility management.

4. Total system integrity must be provided to include insurance against loss of data element records and preparation of appropriate products as described.

5. Costs for hardware and software use, computer time, supplies such as paper for printouts, etc., in support of the preparation of the data element dictionary will be considered a contribution of the computer facility and its parent agency.

ATTACHMENT A

Data Element Description Forms

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Name of Element	(2) Date: 12/3/81

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☒ Yes ☐ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 200 character max.	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

The name of the data element as shown in agency system documentation. Abbreviations are spelled out and qualifiers are added as necessary to fully identify the specific data element.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form; name as given in the specific system, qualifiers added if necessary to clarify the definition.

(18) Output Display Format:

Alphabetized list of these elements, with identification of the source agency and dictionary page on which the element appears, as indexes to the dictionaries.
Also printed in output dictionaries as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

6

DATA ELEME JT DEFINITION FORM

(See GPO Instruction 705.9 'or In truct on on Completion of GPO Form 2443)

(1) Data Element Name	(2) Date
Source Agency	'2-1-8'

(3) System Name:	(4) Standardization Level:			(5) ACTION
DED	☐ GEN	☐ ORG	☒ LOC	☒ ADD

(6) Security Required:	(7) Data Element Use			☐ CHANGE
☐ Yes ☒ No	☐ Input	☐ Output	☒ I O	☐ DELETE

(8) Access Key:	(9) Input Data Element Type:		(10) Input Data Element Length:	(11) Maximum Number of Occurrences:
☒ Yes ☐ No	☐ Numeric	☒ Alphanumeric	4 characters	1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places:	N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

Identifies the agency in whose system the data element appears and which supplied the record for the element.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

A four-character alphabetic acronym to identify the agency source:
DTIC, NASA, NTIS, GPOS, DOES.

(18) Output Display Format:

Carried along in indexes to overall compilation and indexes to separate dictionaries, to identify source of data.
Also printed in output dicitonaries as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

7

GPO Form 2443
(10-81).

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Use	(2) Date: 12-3-81

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Lengthvariable; 12 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

Indicates whether the data element is in one of four possible categories
of use: mandatory, optional, mandatory if available, or required because
of the presence of another cataloging/indexing entry.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

One of four possible entries is selected:
MANDATORY, OPTIONAL, IF AVAILABLE, REQUIRED

(18) Output Display Format:
Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

8

GPO Form 2443
(10-81)

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Definition	(2) Date: 12-3-81

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required. ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☐ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: 720 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

Narrative definition of data element, as recorded in specific agency system documentation.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from system documentation.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

9

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Content		(2) Date: 12-3-81

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 2400 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements.

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/   N/A   2/   3/

(15) Data Element Description:

Indicates what is to be found in this data element in cataloging/ indexing use. Includes rules for recording the data element, special codes to be used if applicable, any limitations imposed by the system such as field length limit, number of occurrences limit, etc.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from system documentation.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

10

GPO Form 2443
(10-81)

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705 9 for In truction on Completion of GPO Form 2443)

(1) Data Element Name: Internal Tag	(2) Date: 2-23-82

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☒ Yes ☐ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 4 character maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places:	N/A

(14) Data Element Synonyms in this System:

1/   N/A   2.   3/

(15) Data Element Description:

The exact content designation by which the data element is identified
for input (often a mnemonic code for the data element).

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Entered as a code; 4 character maximum; left justified with blanks, as
necessary.

(18) Output Display Format:
Four - character code arranged in alpha/numeric sequence, followed by
data element name and source agency iden†ification, as indexes to separate
dictionaries and overall dictionary. Also printed in output dictionaries
as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Internal Tag Description		(2) Date: 2-23-82
(3) System Name: DED	**(4) Standardization Level.** ☐ GEN ☐ ORG ☒ LOC	**(5) ACTION:** ☒ ADD
(6) Security Required: ☐ Yes ☒ No	**(7) Data Element Use:** ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 1200 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/   N/A   2/   3/

(15) Data Element Description:

Any relevant indicators or sub-field codes which accompany the internal tag (especially applicable to GPO data elements).

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from system documentation. If no description is available or relevant, N/A (not applicable) is entered.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

12

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: External Tag		(2) Date: 2-23-82

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☐ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☒ Yes ☐ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: 3 characters	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

The exact content designation by which the data element is identified for exchange of bibliographic data among computer systems.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Entered as 3-character numeric code.

(18) Output Display Format:

Arranged in numeric order, followed by data element name and source agency identification, as indexes to separate dictionaries and to overall dictionary. Also printed in output dictionaries as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

13

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: External Tag Description		(2) Date: 2-23-82

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 1200 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/   N/A   2/   3/

(15) Data Element Description:

Any relevant indicators or sub-field codes which accompany the external tag (especially applicable to GPO data elements).

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from system documentation. If no description is available or relevant, N/A (not applicable) is entered.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705 9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name Character Set		(2) Date 12-3-81
(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 50 character maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements: (12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places:	N/A

(14) Data Element Synonyms in this System

1/   N/A   2/   3/

(15) Data Element Description:

Description of the possible characters that may appear for the given data element in a cataloging/indexing record in the specific agency system.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from the system documentation.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81).

15

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for In truction on Completion of GPO Form 2443)

(1) Data Element Name: Data Source	(2) Date: 12-3-81

(3) System Name: DED	(4) Standardization Level □ GEN □ ORG ☒ LOC	(5) ACTION: ☒ ADD

(6) Security Required: □ Yes ☒ No	(7) Data Element Use. □ Input □ Output ☒ I/O	□ CHANGE □ DELETE

(8) Access Key: □ Yes ☒ No	(9) Input Data Element Type: □ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 8 characters max. per occurrence	(11) Maximum Number of Occurrences: 2

For Numeric Data Elements:

(12) Signed: □ Yes □ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/ N/A   2/   3/

(15) Data Element Description:

Indication as to whether the data that appear in cataloging/indexing records
are taken directly from the source document, are created by the staff
of the agency processing the document, or are created by algorithm or
other automatic means.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Abbreviated version, 8-character maximum, is selected from the list:
DOCUMENT, STAFF, MACHINE, STANDARD.
Up to two occurrences per record are permitted; if two, separate by
semi-colon.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature
of Authorized Person (if CHANGE or DELETE Action)

GPO Form 2443
(10-81)

DATA ELEMENT DEFINITION FORM

(See GPO Instruction 705.9 for Instruction on Completion of GPO Form 2443)

(1) Data Element Name: Notes		(2) Date: 12-3-81

(3) System Name: DED	(4) Standardization Level: ☐ GEN ☐ ORG ☒ LOC	(5) ACTION: ☒ ADD
(6) Security Required: ☐ Yes ☒ No	(7) Data Element Use: ☐ Input ☐ Output ☒ I/O	☐ CHANGE ☐ DELETE

(8) Access Key: ☐ Yes ☒ No	(9) Input Data Element Type: ☐ Numeric ☒ Alphanumeric	(10) Input Data Element Length: variable; 1200 characters maximum	(11) Maximum Number of Occurrences: 1

For Numeric Data Elements:

(12) Signed: ☐ Yes ☐ No	(13) Number of Decimal Places: N/A

(14) Data Element Synonyms in this System:

1/   N/A   2/   3/

(15) Data Element Description:

Any additional relevant information about the data element or its use in the specific system.

(16) Data Element Source:

Specific agency system.

(17) Input Edit Criteria:

Textual form, taken from the system documentation. If notes are not available or relevent, N/A (not applicable) is entered.

(18) Output Display Format:

Printed in output dictionary as recorded.

(19) Data Element Responsibility:

Specific agency and project principal investigator.

Signature _____
of Authorized Person (if CHANGE or DELETE Action)

ATTACHMENT B

Input Data Sources

17a

INPUT DATA SOURCES

Title: Bibliographic data element record.

Medium: Magnetic tape or hardcopy data forms.

Purpose: To prepare a data element dictionary for 5 bibliographic data systems.

Origin: 5 agencies (GPO, NTIS, NASA, DTIC, and DOE).

Frequency: Possibly one-time input, or several increments over a brief
2-weeks to 1-month span.

Volume: From 250 to 500 records, containing several textual fields
of variable length.

Security: Project principal investigator, M.M. Henderson.

Data elements: Name of Element

Source Agency

Use

Definition

Content

Internal Tag

Internal Tag Description

External Tag

External Tag Description

Character Set

Data Source

Notes

All the data elements are defined on GPO Forms 2443 (attached).

Name of Element: Series Statement · Title (Traced)

Source Agency: GPOS.

Use: Mandatory.

Definition: Contains series statements entered under title and is used to
generate series added entries in the same form represented by
the data in the field, exclusive of the International Standard
Serial Number (ISSN), if present.

Content: Contains a series statement (title, volume, or number, etc.) entered
under titles which is used to generate series - added entry.

Internal Tag: SET.

Internal Tag Description: The first indicator is not defined - contains
a blank. The second indicator specifies the number
of characters (0-9) at the beginning of the field
to be ignored in filing.

External Tag: 440.

External Tag Description: N/A.

Character Set: Expanded 8-bit ASCII for MARC records.

Data Source: Document.

Notes: N/A.

Name of Element: Accession Number.

Source Agency: DTIC.

Use: Mandatory.

Definition: A machine processing control number assigned uniquely to a technical report as soon as it has been determined by accessioning procedures that it is not a duplicate report.

Content: A 9-character field composed of 3 alphas followed by 6 numerics. The first two alphas are always AD, the third alpha is a letter code to designate additional intelligence, and the following numerics constitute the sequence number assigned to the document.

Internal Tag: 01.

Internal Tag Description: N/A.

External Tag: 001.

External Tag Description: N/A.

Character Set: Alphanumerics.

Data Source: Staff; Standard.

Notes: A degree of intelligence is structured into the AD number in its initial assignment, to designate classified or unclassified report, and limited or unlimited distribution. Historically, these designations have changed periodically since the inception of AD number assignment. Display of the accession number may include the letter L as the last character, designating a limited distribution report...

ATTACHMENT C

Output Products

OUTPUT PRODUCTS

Title: Data Element Dictionary.

Medium: Paper printout.

Purpose: Examine and compare bibliographic data elements for 5 systems,
toward developing cooperative programs among them.

Frequency: Once.

Volume: 12 copies, 2 for each agency and two for the principal investigator.

Processing: Sort in internal tag number (field number) order, by agency,
then print all data in page format and number pages. Examples
of two possible page formats are attached.

Destination: Two copies to each agency and two copies to the project principal
investigator.

Retention: Until decisions by 5 agencies and principal investigator.

Security: N/A.

Data Elements: Name of Element

                  Source Agency

                  Use

                  Definition

                  Content

                  Internal Tag

                  Internal Tag Description

                  External Tag

                  External Tag Description

                  Character Set

                  Data Source

                  Notes

All the data elements are defined on GPO Forms 2443 (attached).

NAME OF ELEMENT: Series Statement-Title (Traced).

SOURCE AGENCY: GPOS.

USE: Mandatory

DEFINITION: Contains series statements entered under title and is used to
             generate series added entries in the same form represented by
             the data in the field, exclusive of the International Standard
             Serial Number (ISSN), if present.

CONTENT: Contains a series statement (title, volume, or number, etc.) entered
          under title, which is used to generate a series-added entry.

INTERNAL TAG: SET.

INTERNAL TAG DESCRIPTION: The first indicator is not defined - contains a
                           blank. The second indicator specifies the number
                           of characters (0-9) at the beginning of the field
                           to be ignored in filing.

EXTERNAL TAG: 440.

EXTERNAL TAG DESCRIPTION: N/A.

CHARACTER SET: Expanded 8-bit ASCII for MARC records.

DATA SOURCE: Document.

NOTES:

NAME OF ELEMENT: Accession Number.

SOURCE AGENCY: DTIC.

USE: Mandatory.

DEFINITION: A machine processing control number assigned uniquely to a technical
             report as soon as it has been determined by accessioning procedures
             that it is not a duplicate report.

CONTENT: A 9-character field composed of 3 alphas followed by 6 numerics.
          The first two alphas are always AD, the third alpha is a letter
          code to designate additional intelligence, and the following numerics
          constitute the sequence number assigned to the document.

INTERNAL TAG: 01.

INTERNAL TAG DESCRIPTION: N/A.

EXTERNAL TAG: 001.

EXTERNAL TAG DESCRIPTION: N/A.

CHARACTER SET: Alphanumerics.

DATA SOURCE: Staff; Standard.

NOTES: A degree of intelligence is structured into the AD number in its
        initial asignment, to designate classified or unclassified report
        and limited or unlimited distribution. Historically, these designations
        have changed periodically since the inception of AD number assignment.
        Display of the accession number may include the letter L as the 10th
        and last character, designating a limited distribution report, but
        currently this suffix item is machine program-generated from current
        Field 33 (q.v.).

OUTPUT PRODUCTS

Title: Individual and Combined Name of Element Indexes.

Medium: Paper printout.

Purpose: To facilitate examination and comparison of bibliographic data
elements of 5 systems.

Frequency: Once.

Volume: One for each data element dictionary specific to an agency, and
one for the combined dictionary; twelve copies of each.

Processing: Sort in alphabetic order of data element name, add agency
identification and page number for each entry, then list in
sequence. Examples of partial listings for an individual index
and a combined index are attached.

Alternatively and preferably, a KWIC index format would be used,
requiring sorting in alphabetic order of all major words in
the data element name, adding agency ID and page numbers and
listing in sequence as above. Examples for an individual and
a combined index are attached.

Destination: Carried with the dictionaries, in the 12 copies described under
the dictionaries.

Retention: As with the dictionaries.

Security: N/A.

Data Elements: Name of Element

Source Agency

The data elements are defined on GPO Forms 2443 (attached).

INDIVIDUAL AGENCY NAME OF ELEMENT INDEX (alphabetic)

NAME OF ELEMENT	SOURCE AGENCY	PAGE
abstract	DTIC	xx
accession number	DTIC	xx
contract number	DTIC	xx
corporate author	DTIC	xx
descriptive note	DTIC	xx
pagination	DTIC	xx
personal author	DTIC	xx
title	DTIC	xx

COMBINED AGENCIES NAME OF ELEMENT INDEX (alphabetic)

NAME OF ELEMENT	SOURCE AGENCY	PAGE
Abstract	DTIC	xx
Accession number	DTIC	xx
Contract number	DTIC	xx
Corporate author	DTIC	xx
Descriptive note	DTIC	xx
Language code	GPOS	xx
Main entry--personal name	GPOS	xx
Nature of contents code	GPOS	xx
Pagination	DTIC	xx
Personal author	DTIC	xx
Preceding entry	GPOS	xx
Series statement--title (traced)	GPOS	xx
Summary, abstract, scope, etc. note	GPOS	xx
Title	DTIC	xx
Title statement	GPOS	xx

INDIVIDUAL AGENCY NAME OF ELEMENT INDEX (permuted)

NAME OF ELEMENT	SOURCE AGENCY	PAGE
Abstract, scope, etc. note, summary	GPOS	xx
Code, language	GPOS	xx
Code, nature of contents	GPOS	xx
Contents code, nature of	GPOS	xx
Entry--personal name, main	GPOS	xx
Entry, preceding	GPOS	xx
Language code	GPOS	xx
Main entry--personal name	GPOS	xx
Name, main entry--personal	GPOS	xx
Nature of contents code	GPOS	xx
Note, summary, abstract, scope, etc.	GPOS	xx
Personal name, main entry--	GPOS	xx
Preceding entry	GPOS	xx
Scope, etc., note, summary, abstract	GPOS	xx
Series statement--title (traced)	GPOS	xx
Statement, title	GPOS	xx
Statement--title (traced), series	GPOS	xx
Summary, abstract, scope, etc. note	GPOS	xx
Title (traced), series statement--	GPOS	xx
Title statement	GPOS	xx

COMBINED AGENCIES NAME OF ELEMENT INDEX (permuted)

NAME OF ELEMENT	SOURCE AGENCY	PAGE
Abstract	DTIC	xx
Abstract, scope, etc. note, summary	GPOS	xx
Accession number	DTIC	xx
Author, corporate	DTIC	xx
Author, personal	DTIC	xx
Code, language	GPOS	xx
Code, nature of contents	GPOS	xx
Contents code, nature of	GPOS	xx
Contract number	DTIC	xx
Corporate author	DTIC	xx
Descriptive note	DTIC	xx
Entry--personal name, main	GPOS	xx
Entry, preceding	GPOS	xx
Language code	GPOS	xx
Main entry--personal name	GPOS	xx
Name, main entry--personal	GPOS	xx
Nature of contents code	GPOS	xx
Note, descriptive	DTIC	xx
Note, summary, abstract, scope, etc.	GPOS	xx
Number, accession	DTIC	xx
Number, contract	DTIC	xx
Pagination	DTIC	xx
Personal author	DTIC	xx
Personal name, main entry--	GPOS	xx

OUTPUT PRODUCTS

Title: Internal Tag Index.

Medium: Paper printout.

Purpose: To facilitate examination and comparison of bibliographic data
elements of 5 systems.

Frequency: Once.

Volume: One for each data element dictionary specific to an agency and one
for the combined dictionary.

Processing: Sort in order of numbers, letters, and special symbols in initial
3-character portion of tags, add name of element, agency identification
and page number for each entry, then list in sequence of numeric,
alphabetic, special symbol order examples of partial listings
for an individual list and a combined list are attached.

Destination: Carried with the dictionaries, in the 12 copies described
under the dictionaries.

Retention: As with the dictionaries.

Security: N/A.

Data Elements: Name of Element

Source Agency

Internal Tag

The data elements are defined on GPO Forms 2443 (attached).

INDIVIDUAL INTERNAL TAG INDEXES

TAG	NAME OF ELEMENT	SOURCE AGENCY	PAGE
01	Accession number	DTIC	xx
05	Corporate author	DTIC	xx
06	Title	DTIC	xx
09	Descriptive note	DTIC	xx
10	Personal author	DTIC	xx
.	.	.	.
.	.	.	.
.	.	.	.

LAN	Language code	GPOS	xx
MEP	Main entry--personal name	GPOS	xx
NCA	Summary, abstract, scope, etc. note	GPOS	xx
PRE	Preceding entry	GPOS	xx
.	.	.	.
.	.	.	.
.	.	.	.

AUA	Personal author	DOES	xx
CM	Corporate author	DOES	xx
CN	Contract number	DOES	xx
DN	Drop note	DOES	xx
LA	Language	DOES	xx
.	.	.	.
.	.	.	.

COMBINED INTERNAL TAG INDEX

TAG	NAME OF ELEMENT	SOURCE AGENCY	PAGE
.	.	.	.
.	.	.	.
09	Descriptive note	DTIC	xx
10	Personal author	DTIC	xx
.	.	.	.
.	.	.	.
.	.	.	.
AUA	Personal author	DOES	xx
.	.	.	.
.	.	.	.
DN	Drop note	DOES	xx
.	.	.	.
.	.	.	.
MEP	Main entry--personal name	GPOS	xx
NOA	Summary, abstract, scope, etc. note	GPOS	xx

OUTPUT PRODUCTS

<u>Title</u>: External Tag Index.

<u>Medium</u>: Paper printout.

<u>Purpose</u>: To facilitate examination and comparison of bibliographic data.

<u>Frequency</u>: Once. Elements of 5 systems.

<u>Volume</u>: One for each data element dictionary specific to an agency and one for the combined dictionary.

<u>Processing</u>: Sort in numeric order of tag number, add name of element, agency identification, and page number for each entry, then list in sequence. Examples of partial listings for an individual list and a combined list are attached.

<u>Destination</u>: Carried with the dictionaries, in the 12 copies described under the dictionaries.

<u>Retention</u>: As with the dictionaries.

<u>Security</u>: N/A.

<u>Data Elements</u>: Name of Element

Source Agency

External Tag

The data elements are defined on GPO Forms 2443 (attached).

INDIVIDUAL EXTERNAL TAG INDEXES

TAG	NAME OF ELEMENT	SOURCE AGENCY	PAGE
001	Accession number	DTIC	xx
220	Title	DTIC	xx
280	Personal author	DTIC	xx
300	Corporate author	DTIC	xx
320	Contract number	DTIC	xx
520	Pagination	DTIC	xx
540	Descriptive note	DTIC	xx
620	Abstract	DTIC	xx
008	Nature of contents code	GPOS	xx
041	Language code	GPOS	xx
100	Main entry--personal name	GPOS	xx
245	Title statement	GPOS	xx
440	Series statement--title (traced)	GPOS	xx
520	Summary, abstract, scope, etc. note	GPOS	xx
780	Preceding entry	GPOS	xx
010	Accession number	DOES	xx
070	Personal author	DOES	xx
110	Primary title	DOES	xx
240	Contract number	DOES	xx
390	Pagination	DOES	xx
440	Drop note	DOES	xx
710	Corporate author	DOES	xx
950	Abstract	DOES	xx

COMBINED EXTERNAL TAG INDEXES

TAG	NAME OF ELEMENT	SOURCE AGENCY	PAGE
.	.	.	.
.	.	.	.
070	Personal author	DOES	xx
100	Main entry—personal name	GPOS	xx
110	Primary title	DOES	xx
220	Title	DTIC	xx
240	Contract number	DOES	xx
245	Title statement	GPOS	xx
280	Personal author	DTIC	xx
300	Corporate author	DTIC	xx
320	Contract number	DTIC	xx
390	Pagination	DOES	xx
440	Drop note	DOES	xx
440	Series statement—title (traced)	GPOS	xx
520	Summary, abstract, scope, etc. note	GPOS	xx
520	Pagination	DTIC	xx
540	Descriptive note	DTIC	xx
620	Abstract	DTIC	xx
710	Corporate author	DOES	xx
.	.	.	.
.	.	.	.
950	Abstract	DOES	xx